Causes, Clinical Features, and Outcomes from a Prospective Study of Drug-Induced Liver Injury in the United States (DILIN)

Version 2 Prepared by
IMS
3901 Calverton Blvd
Calverton MD 20705
July 2, 2014

Revision History

Version	Author/Title	Date	Comments
1.0	M Spriggs	February 12, 2014	Initial
2.0	M Spriggs	July 2, 2014	V2 Data

Contents

1 Standard Disclaimer

The intent of this DSIC is to provide confidence that the data distributed by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) Repository is a true copy of the study data. Our intent is not to assess the integrity of the statistical analyses reported by study investigators. As with all statistical analyses of complex datasets, complete replication of a set of statistical results should not be expected on a first (or second) exercise in secondary analysis. This occurs for a number of reasons, including differences in the handling of missing data, restrictions on cases included in samples for a particular analysis, software coding used to define complex variables, and other factors. Experience suggests that most discrepancies can ordinarily be resolved by consulting with the study data coordinating center (DCC); however, this process is labor-intensive for both DCC and Repository staff. It is thus not our policy to resolve every discrepancy observed in an integrity check. Specifically, we do not attempt to resolve minor or inconsequential discrepancies with published results or discrepancies that involve complex analyses, *unless NIDDK Repository staff suspect that the observed discrepancy suggests that the dataset may have been corrupted in storage, transmission, or processing by Repository staff*.
We do, however, document in footnotes to the integrity check those instances in which our secondary analyses produced results that were not fully consistent with those reported in the target publication.

2 Study Background

Idiosyncratic drug-induced liver injury (DILI) is among the most common causes of acute liver failure in the United States, accounting for approximately 13% of cases. A prospective study was begun in 2003 to recruit patients with suspected DILI and create a repository of biological samples for analysis. This report summarizes the causes, clinical features, and outcomes from the first 300 patients enrolled.

2.1 Study Methods

The DILIN prospective study is an ongoing multicenter observational study. The study design and procedures were approved by the institutional review board of each clinical center site, and all enrolled patients provided written, fully informed consent. The study design has been described in detail elsewhere.21 In brief, patients (2 years of age or older) were enrolled in this study if there was a strong clinical suspicion that a liver injury event was caused by a medication or an herbal agent occurring within 6 months before enrollment. Additionally, patients must meet one of the following biochemical criteria for enrollment into this study: (1) aspartate aminotransferase (AST) or alanine aminotransferase (ALT) level >5 times the upper limit of normal (ULN) or alkaline phosphatase level >2 times the ULN (or pretreatment baseline if baseline level is abnormal) on 2 consecutive occasions, (2) total serum bilirubin level >2.5 mg/dL along with elevated AST or ALT or alkaline phosphatase level, or (3) international normalized ratio (INR) >1.5 with elevated AST or ALT or alkaline phosphatase level. Known or suspected acetaminophen toxicity and a history of bone marrow or liver transplantation before the liver injury event were exclusion criteria. Patients with underlying hepatitis C virus (HCV), hepatitis B virus, or nonalcoholic fatty liver disease were eligible if they developed superimposed DILI; however, those with other types of underlying chronic liver disease (eg, autoimmune liver disease, sclerosing cholangitis) were ineligible. Subjects with known or suspected acetaminophen hepatotoxicity or a history of bone marrow or liver transplantation were excluded.

3 Archived Datasets

The DCC submitted 102 datasets that were used for the analysis for this publication. Note that a separate file was sent to limit the records to the first 300 subjects enrolled. We used 13 datasets for this DSIC: JUDGMNT, HXINJ2, LABFLOW, DEMOG, LIVHX, QUESTS, MEDHX1, MEDHX2, HXINJ2, PEX1, DEIDMAN1, CHRONIC, PTCOMPL. Separately, the variable STDLABHX was used to confirm that the Eosinophils counts in the tables were not reproducible.

4 Statistical Methods

We compared our DSIC results to the published results in:

- Table 1. Characteristics of Subjects With Different Patterns of Liver Injury
- Table 2. Clinical Characteristics of DILI Subjects Categorized by Liver Injury Severity

Our DSIC analyses were conducted in SAS v9 (Appendix 1). The SAS code and output used to support the findings of the DSIC appear as Appendix 1.

Characteristics of Subjects With Different Patterns of Liver Injury is in Table 1, which presents study Ns and percentages as well as means ± standard deviations where appropriate. Clinical Characteristics of DILI Subjects Categorized by Liver Injury Severity is in Table 2, which presents study Ns and percentages as well as means ± standard deviations, medians and interquartile ranges where appropriate.

5 Results

Variables used to replicate Table 1. **Characteristics of Subjects With Different Patterns of Liver Injury** are shown in Table A.

Table A: Variables Used to Replicate Table 1.

Measure	Variable
Age, mean ± SD (y)	Not Reproducible. Demog.AGECAT is categorical.
Proportion aged 65 years or older (%)	Not Reproducible. Demog.AGECAT is categorical.
Female (%)	Demog.SEX
Self-reported race (%)	Demog.WHITE demog.BLACK demog.ASIAN
Body mass index, mean ± SD (kg/m2)	pex1.HT, pex1.WT, pex1.HTUN, pex1.WTUN
Alcohol use (%)	Quests.ALCOHOL
Preexisting liver disease (%)	Livhx.LIVER_DISEASE
Diabetes mellitus (%)	MEDHX1.DIABETES
Liver biochemistries, DILI recognition (mean ± SD)	LABFLOW.ALTVAL, LABFLOW.AKPVAL, LABFLOW.STBVAL, LABFLOW.INRRATE, LABFLOW.FLWRATIO, LABFLOW.FCLABDT, hxinj2.ONSETDT
Liver biochemistries, peak values (mean ± SD)	LABFLOW.ALTVAL, LABFLOW.AKPVAL, LABFLOW.STBVAL, LABFLOW.INRRATE
Absolute eosinophils/µL (mean ± SD)	Not Reproducible: Eosinophils are only available as a percentage in stdlabhx.LABHXN06,LABHXV06 and LABHXU06.
Severity of liver injury (%)	Judgmnt.SEVERITY
Causality assessment (%)	Judgmnt.ASSESS

DSIC Results: Table 1. The published manuscript results and the DSIC results for Table 1 are shown on the next pages (Table B). The base Ns, means and standard deviations for the patient characteristics and histology results calculated by the DSIC generally correspond to published values with expected differences because the data sent is not the exact same data used in the paper. The current data includes data that was taken after the publication. All differences are noted below and results that were not reproducible are marked n/a.

Table B: Table 1. Characteristics of Subjects With Different Patterns of Liver Injury.

	Chalasani *et al (2008)*		DSIC	
	Entire Cohort (n=300)	Hepatocellular (n=169)	Entire Cohort (n=300)	Hepatocellular (n=171)
Age, mean ± SD (y)	48 ± 18	44 ± 18	n/a	n/a
Proportion aged 65 years or older (%)	18.5	12	n/a	n/a
Female (%)	60	65	60	65
Self-reported race (%)				
White	79	76	79	75
Black	11	11	11	12
Asian	3.7	4.7	4	5
Body mass index, mean ± SD (kg/m2)	26.8 ± 6.5	27 ± 7.4	26.8 ± 6.5	26.7 ± 7.3
Alcohol use (%)	51	50	51	49
Preexisting liver disease (%)	5.7	8	16	14
Diabetes mellitus (%)	27	16	32	28
Liver biochemistries, DILI recognition (mean ± SD)				
ALT (U/L)	788 ± 967	1157 ± 1131	787 ± 959	1166 ± 1115
Alkaline phosphatase (U/L)	295 ± 272	190.5 ± 119.5	291 ± 264	192 ± 123
Total bilirubin (mg/dL)	6.3 ± 6.3	6.2 ± 7.1	6.3 ± 6.5	6.1 ± 7.0
INR	1.5 ± 0.9	1.6 ± 1.1	1.4 ± 0.9	1.6 ± 1.1
Liver biochemistries, peak values (mean ± SD)				
ALT (U/L)	985 ± 1168	1426 ± 1314	1015 ± 1189	1484 ± 1363
Alkaline phosphatase (U/L)	390 ± 382	248 ± 167	397 ± 390	256 ± 186
Total bilirubin (mg/dL)	11.4 ± 10.2	10.5 ± 10.1	11.8 ± 10.7	11.0 ± 10.6
INR	1.6 ± 1.4	1.9 ± 1.9	1.7 ± 1.6	1.9 ± 2.0
Absolute eosinophils/µL (mean ± SD)	210 ± 310	157 ± 153	n/a	n/a
Severity of liver injury (%)				
Mild	27	33	25	30
Moderate	19	15	19	14
Moderate-hospitalized	33	27	27	24
Severe	15	16	20	22
Fatal	6	9	9	11
Causality assessment (%)				
Definite	32	31	31	33
Very likely	41	43	41	41
Probable	14	13.4	14	15
Possible	20	10	10	9
Unlikely	3	2.8	3	2

	Chalasani *et al (2008)*		DSIC	
	Cholestatic (n=68)	Mixed (n=61)	Cholestatic (n=67)	Mixed (n=61)
Age, mean ± SD (y)	54 ± 16	54 ± 18	n/a	n/a
Proportion aged 65 years or older (%)	26.5	26	n/a	n/a
Female (%)	50	57	51	56
Self-reported race (%)				
White	84	85	84	84
Black	9	12	9	11
Asian	1.5	1	1	2
Body mass index, mean ± SD (kg/m2)	27 ± 5.8	27 ± 4.7	26.8 ± 5.8	27.1 ± 4.7
Alcohol use (%)	50	57	49	57
Preexisting liver disease (%)	7	0	29	0
Diabetes mellitus (%)	25	23	40	36
Liver biochemistries, DILI recognition (mean± SD)				
ALT (U/L)	203 ± 160	384 ± 206	192 ± 159	378 ± 208
Alkaline phosphatase (U/L)	532 ± 386	305 ± 236	530 ± 382	304 ± 228
Total bilirubin (mg/dL)	7.3 ± 5.6	5.5 ± 4.6	7.6 ± 6.4	5.2 ± 4.6
INR	1.2 ± 0.3	1.3 ± 0.4	1.2 ± 0.3	1.2 ± 0.4
Liver biochemistries, peak values (mean ± SD)				
ALT (U/L)	314 ± 451	465 ± 295	308 ± 422	480 ± 301
Alkaline phosphatase (U/L)	703 ± 542	378 ± 323	742 ± 553	413 ± 361
Total bilirubin (mg/dL)	14 ± 10.8	10.2 ± 9.4	14.4 ± 11.5	10.8 ± 9.7
INR	1.3 ± 0.5	1.3 ± 0.5	1.5 ± 1.0	1.4 ± 0.9
Absolute eosinophils/µL (mean ± SD)	221 ± 207	389 ± 653	n/a	n/a
Severity of liver injury (%)				
Mild	17	23	16	23
Moderate	20	29	24	30
Moderate-hospitalized	44	36.5	28	33
Severe	17	10	24	11
Fatal	2	2	7	3
Causality assessment (%)				
Definite	27	38	22	36
Very likely	34	43	40	46
Probable	22	8	19	8
Possible	12	8	13	7
Unlikely	5.1	4	4	3

Variables used to replicate Table 2. **Clinical Characteristics of DILI Subjects Categorized by Liver Injury Severity** are shown in Table C.

Table C: Variables Used to Replicate Table 2.

Measure	Variable
Age, mean ± SD (y)x	Not Reproducible. Demog.AGECAT is categorical.
Proportion aged 65 years or older (%)	Not Reproducible. Demog.AGECAT is categorical.
Female (%)	Demog.SEX
Self-reported race (%)	Demog.WHITE demog.BLACK demog.ASIAN
Body mass index, mean ± SD (kg/m2)	pex1.HT, pex1.WT, pex1.HTUN, pex1.WTUN
Alcohol use (%)	Quests.ALCOHOL
Preexisting liver disease (%)	Livhx.MEDLIVHX
Diabetes mellitus (%)	MEDHX1.PASTHX
Days between exposure and DILI recognition, median (25^{th}, 75^{th} percentiles)	Not Reproducible: The calculated values from hxinj2.ONSETDT and labflow.FCLABDT negatives and do not match these numbers.
Implicated agent(s) (%)	Unduplicateable: Not enough information about how these were categorized exists.
Liver biochemistries, values at DILI onset (mean ± SD)	LABFLOW.ALTVAL, LABFLOW.AKPVAL, LABFLOW.STBVAL, LABFLOW.INRRATE, LABFLOW.FLWRATIO, LABFLOW.FCLABDT, hxinj2.ONSETDT
Liver biochemistries, peak values (mean ± SD)	LABFLOW.ALTVAL, LABFLOW.AKPVAL, LABFLOW.STBVAL, LABFLOW.INRRATE
Absolute eosinophils/µL (mean ± SD)	Not Reproducible: Eosinophils are only available as a percentage in stdlabhx.LABHXN06,LABHXV06 and LABHXU06.
Pattern of liver injury (%)	Judgmnt.SEVERITY

Measure	Variable
Causality assessment (%)	Judgmnt.ASSESS
Chronic DILI %	Chronic.CINEXC1-CINEXC10
Liver-related mortality (%)	Ptcompl.NLLRDTH
Liver transplantation (%)	hxinj2.LIVTRX

DSIC Results: Table 2. The published manuscript results and the DSIC results for Table 2 are shown on the next pages (Table D). The base Ns, means and standard deviations for the patient characteristics and histology results calculated by the DSIC generally correspond to published values with expected differences because the data sent is not the exact same data used in the paper. The current data also includes data that was taken after the published. All differences are noted below and results that were not reproducible are marked n/a.

Table D: Table 2. Clinical Characteristics of DILI Subjects Categorized by Liver Injury Severity.

	Chalasani *et al (2008)*		DSIC	
	Mild/moderate DILI (n = 195)	Severe DILI (n = 51)	Mild/moderate DILI (n = 213)	Severe DILI (n = 86)
Age, mean ± SD (y)	50 ± 18	46 ± 16.3	n/a	n/a
Proportion aged ≥65 years (%)	23	11	n/a	n/a
Female (%)	57	59	59	63
Self-reported race (%)				
White	81	67	83	69
Black	11	18	10	14
Asian	2	8	2	8
Body mass index, mean ± SD (kg/m2)	26.4 ± 5.9	27.2 ± 7.8	26.6 ± 6.1	27.4 ± 7.5
Alcohol use (%)	57	32	56	39
Preexisting liver disease (%)	6	9	16	17
Diabetes (%)	25	37	26	47
Days between exposure and DILI recognition, median (25th, 75th percentiles)	35.5 (19, 89)	65.5 (33, 263)	n/a	n/a
Implicated agent(s) (%)				
Single prescription agent	74	80	n/a	n/a
Single or multiple dietary supplements	7.2	2	n/a	n/a
Multiple prescription or prescription plus dietary supplements	19	18	n/a	n/a
Liver biochemistries, values at DILI onset (mean ± SD)				
ALT (U/L)	610 ± 685	1218 ± 1559	658 ± 710	1112 ± 1349
Alkaline phosphatase (U/L)	309 ± 284	283 ± 247	301 ± 281	268 ± 218
Bilirubin (mg/dL)	5.2 ± 5.1	9.1 ± 8.4	4.9 ± 5.0	9.7 ± 8.3
INR	1.2 ± 0.3	2.3 ± 1.6	1.1 ± 0.3	1.9 ± 1.3
Liver biochemistries, peak values (mean ± SD)				
ALT (U/L)	733 ± 726	1513 ± 1734	816 ± 839	1516 ± 1688
Alkaline phosphatase (U/L)	388 ± 399	401 ± 354	375 ± 352	453 ± 469
Bilirubin (mg/dL)	8.9 ± 9.4	18.4 ± 10.2	8.7 ± 8.9	19.5 ± 11.0
INR	1.2 ± 0.3	2.9 ± 2.5	1.2 ± 0.5	2.9 ± 2.5
Absolute eosinophil count/µL (mean ± SD)	195 ± 344	197 ± 293	n/a	n/a

	Chalasani *et al (2008)*		DSIC	
	Mild/moderate DILI (n = 195)	Severe DILI (n = 51)	Mild/moderate DILI (n = 213)	Severe DILI (n = 86)
Pattern of liver injury (%)				
Hepatocellular	50	69	54	65
Cholestatic	27	19	22	24
Mixed	23	12.5	25	10
Causality score				
Definite	32	25.5	34	24
Highly likely	41.5	35	42	40
Probable	13	22	13	19
Possible	10	10	9	13
Unlikely	2.6	8	2	5
Chronic DILI (%)	17	12	19	30
Liver-related mortality (%)	0	23	0	10
Liver transplantation (%)	0	2.9	0	9

6 Conclusions

The results of these DSIC analyses provide confidence that the DILIN data distributed by the NIDDK repository are a true representation of the study data.

7 References

Chalasani N, Fontana RJ, Bonkovsky HL, Watkins PB, Davern T, Serrano J, et al. Causes, clinical features, and outcomes from a prospective study of drug-induced liver injury in the United States. Gastroenterology 2008; 135: 1924-1934, 1934 e1-e4.

Appendix 1. SAS Output used to Replicate Manuscript Results.

```
title1 "%sysfunc(getoption(sysin))";
title2 " ";

options nofmterr source2 mprint ls=155;
%include "/prj/niddk/ims_analysis/DILIN_PRO/private_created_data/pro_formats.edit.sas";
libname sas_out "/prj/niddk/ims_analysis/DILIN_PRO/prog_initial_analysis/";
libname fin_pop "/prj/niddk/ims_analysis/DILIN_PRO/private_orig_data/dilin_pro_deidentify_grant02_draft";
libname fin3pop "/prj/niddk/ims_analysis/DILIN_PRO/private_orig_data/dilin_pro_final_deidentify_grant01_manuscript01_pop_final";

data fin3pop;
  set fin3pop.deidman1;

data accumfreq1 accumfreq3 accummeans3 accummedian accumfreq2 accummeans2;
  set _null_;

*** Macro ***;

%macro idmerger(dsn=);

data &dsn._300v;
  merge fin_pop.&dsn.(in=in_dsn)
        fin3pop(in=in_man);
  by DEIDNUM;
  if in_dsn and in_man then output;
%mend;

%macro freqdata(invar=);
proc freq data=analysis compress noprint;
     tables &invar/out=data1;
     format _all_;
     run;

data data1(keep=flow_num LEVEL name CHARALL);
  set data1(rename=(&invar=LEVEL));
  length name $100 CHARALL $100;
  name=upcase("&invar");
  PCT_DISP=round(PERCENT);
  flow_num=.A;
  CHARALL=compress(put(PCT_DISP,8.));

data accumfreq1;
  set accumfreq1 data1;
%mend freqdata;
%macro freqdata2(invar=);
```

```
proc freq data=analysis compress noprint;
     tables flow_num*&invar/out=data1 outpct;
     format _all_;
     run;

data data1(keep=flow_num LEVEL name CHARALL);
  set data1(rename=(&invar=LEVEL));
  length name $100;
  name=upcase("&invar");
  PCT_DISP=round(PCT_ROW);
  CHARALL=compress(put(PCT_DISP,8.));

data accumfreq2;
  set accumfreq2 data1;
%mend freqdata2;

%macro freqdata3(invar=);
proc freq data=analysis compress noprint;
     tables JUDGE_FLAG*&invar/out=data1 outpct;
     format _all_;
     run;

data data1(keep=JUDGE_FLAG LEVEL name CHARALL);
  set data1(rename=(&invar=LEVEL));
  length name CHARALL $100 ;
  name=upcase("&invar");
  PCT_DISP=round(PCT_ROW);
  CHARALL=compress(put(PCT_DISP,8.));

data accumfreq3;
  set accumfreq3 data1;
%mend freqdata3;

%macro meandata2(invar=, roundvar=, digit=);
proc means data=analysis mean stddev noprint;
     var &invar;
     class flow_num;
     output out=data1 mean=mean stddev=stddev;
     run;

data data1(drop=_TYPE_ mean stddev rename=(_FREQ_=COUNT));
  set data1;
  length name CHARALL $100;
  name=upcase("&invar");
  mean=round(mean,&roundvar);
  stddev=round(stddev,&roundvar);
  CHARALL=compress(put(mean,8.&digit))||" ± "||compress(put(stddev,8.&digit));
  if flow_num=. then flow_num=.A;
```

```
data accummeans2;
  set accummeans2 data1;

%mend meandata2;

%macro meandata3(invar=, roundvar=, digit=);
proc means data=analysis mean stddev noprint;
     var &invar;
     class JUDGE_FLAG;
     output out=data1 mean=mean stddev=stddev;
     run;

data data1(drop=_TYPE_ mean stddev rename=(_FREQ_=COUNT));
  set data1;
  length name CHARALL $100;
  if JUDGE_FLAG=. then delete;
  name=upcase("&invar");
  mean=round(mean,&roundvar);
  stddev=round(stddev,&roundvar);
  CHARALL=compress(put(mean,8.&digit))||" ± "||compress(put(stddev,8.&digit));


data accummeans3;
  set accummeans3 data1;

%mend meandata3;

%macro mediandata(invar=, roundvar=, digit=);
proc means data=analysis median p25 p75;
     var &invar;
     class JUDGE_FLAG;
     output out=data1 median=median p25=p25 p75=p75;
     title3 ' ';
     run;

data data1(drop= median p25 p75);
  set data1;
  length name CHARALL $100;
  if JUDGE_FLAG=. then delete;
  name=upcase("&invar");
  median=round(median,&roundvar);
  CHARALL=compress(put(median,8.&digit))||" ( "||compress(put(p25,8.))||", "||compress(put(p75,8.))||")";


data accummedian;
  set accummedian data1;
```

```
%mend mediandata;

%macro inert(orderer=);
  orderer=&orderer;
  flow_num=.A;
  output;
  flow_num=1;
  output;
  flow_num=2;
  output;
  flow_num=3;
  output;
%mend inert;

%macro inert2(orderer=);
  orderer=&orderer;
  JUDGE_FLAG=0;
  output;
  JUDGE_FLAG=1;
  output;
%mend inert2;

%idmerger(dsn=judgmnt);
%idmerger(dsn=hxinj2);
%idmerger(dsn=labflow);
%idmerger(dsn=demog);
%idmerger(dsn=livhx);
%idmerger(dsn=quests);
%idmerger(dsn=medhx1);
%idmerger(dsn=medhx2);
%idmerger(dsn=hxinj2);
%idmerger(dsn=pex1);
%idmerger(dsn=chronic);
%idmerger(dsn=ptcompl);

*** used to show Eosinophils problem only ***;
%idmerger(dsn=stdlabhx);

proc format;
  value fratio
  5-High="HEPATOCELLULAR"
  0-2="CHOLESTATIC"
  2<-<5="MIXED"
  ;
  value fnratio
  .A="ENTIRE COHORT"
  1="HEPATOCELLULAR"
  2="CHOLESTATIC"
```

```
  3="MIXED"
  ;
  value judge
  0="Mild/Moderate"
  1="Severe"
  ;

*** Dataset construction ***;

*** judgment: SEVERITY, ASSESS, JUDGE_FLAG***;
*** FORM: keep reassessments when there are multiples ***;
data judgment(keep=DEIDNUM FORM SEVERITY ASSESS JUDGE_FLAG);
  set judgmnt_300v;
  if SEVERITY in(1 2 3) then JUDGE_FLAG=0;
  else if SEVERITY in(4 5) then JUDGE_FLAG=1;
  else abort;
  if REVIEW=4 and strip(FORM)="INITIAL_ASSESSMENT" then output;

proc sort data=judgment nodupkeys;
  by _ALL_;

proc sort data=judgment;
  by DEIDNUM;

data judgment;
  set judgment;
  by DEIDNUM;
  if last.DEIDNUM=1 then output;

*** hx_lab: ALTVAL AKPVAL STBVAL INRRATE flow_num;

data labflow_300v_nomiss;
  set labflow_300v;
  if FLWRATIO=. then delete;

proc sort data=labflow_300v_nomiss;
  by DEIDNUM FCLABDT;

proc sort data=hxinj2_300v;
  by DEIDNUM;

data hx_lab_prelim;
  merge hxinj2_300v(in=in_hx keep=DEIDNUM ONSETDT) labflow_300v_nomiss(in=in_lab);
  by DEIDNUM;
  hx_ok=in_hx;
  lab_ok=in_lab;
  retain OUTPUT_DATE OUTPUT_FLAG EARLY_TO_ONSET;
  if first.DEIDNUM then do;
```

```
    OUTPUT_FLAG=0;
    OUTPUT_DATE=.;
    EARLY_TO_ONSET=ONSETDT-FCLABDT;
  end;
  if FCLABDT>=ONSETDT and OUTPUT_FLAG=0 then do;
    OUTPUT_DATE=FCLABDT;
    OUTPUT_FLAG=1;
  end;
  if OUTPUT_FLAG=1 and FCLABDT ne OUTPUT_DATE then do;
    OUTPUT_FLAG=2;
    OUTPUT_DATE=.;
  end;

data hx_lab_prelim_b;
  set hx_lab_prelim;
  if OUTPUT_FLAG=1 then output;

proc sort data=hx_lab_prelim_b;
  by DEIDNUM FLWRATIO;

data hx_lab(keep=DEIDNUM ALTVAL AKPVAL STBVAL INRRATE FLWRATIO EARLY_TO_ONSET);
  set hx_lab_prelim_b;
  by DEIDNUM;
  if last.DEIDNUM then output;



data hx_lab(drop=FLWRATIO);
  set hx_lab;
  if FLWRATIO>=5 then FLOW_NUM=1;
  else if 0<=FLWRATIO<=2 then FLOW_NUM=2;
  else if 2<FLWRATIO<5 then FLOW_NUM=3;

*** Demog: demog_300v, SEX WHITE BLACK ASIAN RACEOTHR AGECAT ***;
data demog_300v(keep=DEIDNUM SEX WHITE BLACK ASIAN RACEOTHR AGECAT);
  set demog_300v;
  by DEIDNUM;
  if WHITE=. then WHITE=0;
  if BLACK=. then BLACK=0;
  if ASIAN=. then ASIAN=0;

*** Liver Disease: livhx_300v_s, LIVER_DISEASE ***;

data livhx_300v_s(keep=DEIDNUM LIVER_DISEASE);
  set livhx_300v;
  by DEIDNUM;
  length LIVER_DISEASE 8.;
  retain LIVER_DISEASE;
  if first.DEIDNUM then LIVER_DISEASE=0;
```

```
  if MEDLIVHX in(3 4) then LIVER_DISEASE=1;
  if last.DEIDNUM then output;

*** Lab Flow: labflow_300v_s, ALT_PEAK,ALK_PEAK,BILI_PEAK, INR_PEAK ***;

data labflow_300v_s(keep=DEIDNUM ALT_PEAK ALK_PEAK BILI_PEAK INR_PEAK);
  set LABFLOW_300v;
  by DEIDNUM;
  length ALT_PEAK ALK_PEAK BILI_PEAK INR_PEAK 8.;
  retain ALT_PEAK ALK_PEAK BILI_PEAK INR_PEAK;
  if first.DEIDNUM then do;
    ALT_PEAK=.;
    ALK_PEAK=.;
    BILI_PEAK=.;
    INR_PEAK=.;
  end;
  ALT_PEAK=max(ALTVAL,ALT_PEAK);
  ALK_PEAK=max(AKPVAL,ALK_PEAK);
  BILI_PEAK=max(STBVAL,BILI_PEAK);
  INR_PEAK=max(INRRATE,INR_PEAK);

  if last.DEIDNUM then output;

***  quests_300v, ALCOHOL ***;
data quests_300v(keep=DEIDNUM ALCOHOL);
  set quests_300v;

***  MEDHX1_300V_s, DIABETES ***;
data MEDHX1_300V_s;
  set MEDHX1_300v;
  by DEIDNUM;
  length DIABETES 8.;
  retain DIABETES;
  if first.DEIDNUM then DIABETES=0;
  if PASTHX in(1) then DIABETES=1;
  if last.DEIDNUM then output;

***  MEDHX2_300v, LIVHXYN ***;
data MEDHX2_300v(keep=DEIDNUM LIVHXYN);
  set MEDHX2_300v;
  by DEIDNUM;

***  hxinj2_300v, LIVER_TRANSPLANT ***;
data hxinj2_300v(keep=DEIDNUM LIVER_TRANSPLANT);
  set hxinj2_300v;
  by DEIDNUM;
  length LIVER_TRANSPLANT 8.;
  retain LIVER_TRANSPLANT;
```

```
  if first.DEIDNUM then LIVER_TRANSPLANT=0;
  if LIVTRX in(1) then LIVER_TRANSPLANT=1;
  if last.DEIDNUM then output;

*** pex1, BMI ***;

data pex1_300v(keep=DEIDNUM BMI);
  set pex1_300v;
  by DEIDNUM;
  if htun=1 then bmi_height=ht/100;
  else if htun=2 then bmi_height=ht/39.3701;
  if wtun=1 then bmi_weight=wt;
  else if wtun=2 then bmi_weight=wt/2.20462;
  if bmi_height>. and bmi_weight>. then bmi=bmi_weight/(bmi_height*bmi_height);
  if first.DEIDNUM then output;

*** chronic, CHRONIC ***;

data chronic_300v(keep=DEIDNUM CHRONIC);
  set chronic_300v;
  if CINEXC1=1 or CINEXC2=1 or CINEXC3=1 or CINEXC4=1 or CINEXC5=1 or CINEXC6=1 or CINEXC7=1 or CINEXC8=1 or CINEXC9=1 or CINEXC10=1 then CHRONIC=1;
  else CHRONIC=0;

*** ptcompl, NLLRDTH ***;

data ptcompl_300v(keep=DEIDNUM NLLRDTH);
  set ptcompl_300v;
  if NLLRDTH=. then NLLRDTH=0;

*** data combination ****;
data analysis;
  merge judgment hx_lab demog_300v livhx_300v_s labflow_300v_s quests_300v MEDHX1_300V_s MEDHX2_300v hxinj2_300v pex1_300v chronic_300v ptcompl_300v;
  by DEIDNUM;

%freqdata(invar=SEX);
%freqdata(invar=WHITE);
%freqdata(invar=BLACK);
%freqdata(invar=ASIAN);
%freqdata(invar=ALCOHOL);
%freqdata(invar=LIVER_DISEASE);
%freqdata(invar=DIABETES);
%freqdata(invar=SEVERITY);
%freqdata(invar=ASSESS);
%freqdata2(invar=SEX);
%freqdata2(invar=WHITE);
%freqdata2(invar=BLACK);
%freqdata2(invar=ASIAN);
%freqdata2(invar=ALCOHOL);
```

```
%freqdata2(invar=LIVER_DISEASE);
%freqdata2(invar=DIABETES);
%freqdata2(invar=SEVERITY);
%freqdata2(invar=ASSESS);
%meandata2(invar=BMI, roundvar=.1, digit=1);
%meandata2(invar=ALTVAL, roundvar=1, digit=0);
%meandata2(invar=AKPVAL, roundvar=1, digit=0);
%meandata2(invar=STBVAL, roundvar=.1, digit=1);
%meandata2(invar=INRRATE, roundvar=.1, digit=1);
%meandata2(invar=ALT_PEAK, roundvar=1, digit=0);
%meandata2(invar=ALK_PEAK, roundvar=1, digit=0);
%meandata2(invar=BILI_PEAK, roundvar=.1, digit=1);
%meandata2(invar=INR_PEAK, roundvar=.1, digit=1);

data accumfreq;
  set accumfreq1 accumfreq2;

data accummeans;
  set accummeans2(drop=COUNT);

data accumfreqmeans;
  set accumfreq accummeans;


data accuminert;
  retain orderer flow_num;
  %inert(orderer=1);
  %inert(orderer=2);
  %inert(orderer=4);
  %inert(orderer=12);
  %inert(orderer=17);
  %inert(orderer=23);
  %inert(orderer=24);
  %inert(orderer=29);


data accumall;
  set accumfreqmeans accuminert;
  if FLOW_NUM=. then delete;
  if name="SEX" and LEVEL=1 then delete;
  if name="SEX" and LEVEL=2 then orderer=3;
  if name="WHITE" and LEVEL=0 then delete;
  if name="WHITE" and LEVEL=1 then orderer=5;
  if name="BLACK" and LEVEL=0 then delete;
  if name="BLACK" and LEVEL=1 then orderer=6;
  if name="ASIAN" and LEVEL=0 then delete;
  if name="ASIAN" and LEVEL=1 then orderer=7;
  if name="BMI" then orderer=8;
```

```
  if name="ALCOHOL" and LEVEL=. then delete;
  if name="ALCOHOL" and LEVEL=0 then delete;
  if name="ALCOHOL" and LEVEL=1 then orderer=9;
  if name="LIVER_DISEASE" and LEVEL=. then delete;
  if name="LIVER_DISEASE" and LEVEL=0 then delete;
  if name="LIVER_DISEASE" and LEVEL=1 then orderer=10;
  if name="DIABETES" and LEVEL=. then delete;
  if name="DIABETES" and LEVEL=0 then delete;
  if name="DIABETES" and LEVEL=1 then orderer=11;
  if name="ALTVAL" then orderer=13;
  if name="AKPVAL" then orderer=14;
  if name="STBVAL" then orderer=15;
  if name="INRRATE" then orderer=16;
  if name="ALT_PEAK" then orderer=18;
  if name="ALK_PEAK" then orderer=19;
  if name="BILI_PEAK" then orderer=20;
  if name="INR_PEAK" then orderer=21;
  if name="SEVERITY" and level=. then delete;
  if name="SEVERITY" and level=1 then orderer=25;
  if name="SEVERITY" and level=2 then orderer=26;
  if name="SEVERITY" and level=3 then orderer=27;
  if name="SEVERITY" and level=4 then orderer=28;
  if name="SEVERITY" and level=5 then orderer=29;
  if name="ASSESS" and level=. then delete;
  if name="ASSESS" and level=1 then orderer=31;
  if name="ASSESS" and level=2 then orderer=32;
  if name="ASSESS" and level=3 then orderer=33;
  if name="ASSESS" and level=4 then orderer=34;
  if name="ASSESS" and level=5 then orderer=35;

proc sort data=accumall;
  by flow_num orderer;


*** Table2 ***;

%freqdata3(invar=SEX);
%freqdata3(invar=WHITE);
%freqdata3(invar=BLACK);
%freqdata3(invar=ASIAN);
%freqdata3(invar=ALCOHOL);
%freqdata3(invar=LIVER_DISEASE);
%freqdata3(invar=DIABETES);
%freqdata3(invar=ASSESS);
%freqdata3(invar=LIVER_TRANSPLANT);
%freqdata3(invar=FLOW_NUM);
%freqdata3(invar=CHRONIC);
%freqdata3(invar=NLLRDTH);
```

```
%meandata3(invar=BMI, roundvar=.1, digit=1);
%meandata3(invar=ALTVAL, roundvar=1, digit=0);
%meandata3(invar=AKPVAL, roundvar=1, digit=0);
%meandata3(invar=STBVAL, roundvar=.1, digit=1);
%meandata3(invar=INRRATE, roundvar=.1, digit=1);
%meandata3(invar=ALT_PEAK, roundvar=1, digit=0);
%meandata3(invar=ALK_PEAK, roundvar=1, digit=0);
%meandata3(invar=BILI_PEAK, roundvar=.1, digit=1);
%meandata3(invar=INR_PEAK, roundvar=.1, digit=1);


data accuminert;
  retain orderer JUDGE_FLAG;
  %inert2(orderer=1);
  %inert2(orderer=2);
  %inert2(orderer=4);
  %inert2(orderer=12);
  %inert2(orderer=13);
  %inert2(orderer=14);
  %inert2(orderer=15);
  %inert2(orderer=16);
  %inert2(orderer=17);
  %inert2(orderer=22);
  %inert2(orderer=27);
  %inert2(orderer=32);

data accumall2;
  set accumfreq3 accummeans3 accummedian accuminert;
  if JUDGE_FLAG=. then delete;
  if name="SEX" and LEVEL=1 then delete;
  if name="SEX" and LEVEL=2 then orderer=3;
  if name="WHITE" and LEVEL=0 then delete;
  if name="WHITE" and LEVEL=1 then orderer=5;
  if name="BLACK" and LEVEL=0 then delete;
  if name="BLACK" and LEVEL=1 then orderer=6;
  if name="ASIAN" and LEVEL=0 then delete;
  if name="ASIAN" and LEVEL=1 then orderer=7;
  if name="BMI" then orderer=8;
  if name="ALCOHOL" and LEVEL=. then delete;
  if name="ALCOHOL" and LEVEL=0 then delete;
  if name="ALCOHOL" and LEVEL=1 then orderer=9;
  if name="LIVER_DISEASE" and LEVEL=. then delete;
  if name="LIVER_DISEASE" and LEVEL=0 then delete;
  if name="LIVER_DISEASE" and LEVEL=1 then orderer=10;
  if name="DIABETES" and LEVEL=. then delete;
  if name="DIABETES" and LEVEL=0 then delete;
  if name="DIABETES" and LEVEL=1 then orderer=11;
  if name="ALTVAL" then orderer=18;
```

```
  if name="AKPVAL" then orderer=19;
  if name="STBVAL" then orderer=20;
  if name="INRRATE" then orderer=21;
  if name="ALT_PEAK" then orderer=23;
  if name="ALK_PEAK" then orderer=24;
  if name="BILI_PEAK" then orderer=25;
  if name="INR_PEAK" then orderer=26;
  if name="FLOW_NUM" and level=. then delete;
  if name="FLOW_NUM" and level=1 then orderer=29;
  if name="FLOW_NUM" and level=2 then orderer=30;
  if name="FLOW_NUM" and level=3 then orderer=31;
  if name="ASSESS" and level=. then delete;
  if name="ASSESS" and level=1 then orderer=33;
  if name="ASSESS" and level=2 then orderer=34;
  if name="ASSESS" and level=3 then orderer=35;
  if name="ASSESS" and level=4 then orderer=36;
  if name="ASSESS" and level=5 then orderer=37;
  if name="CHRONIC" and level in(0 .) then delete;
  if name="CHRONIC" and level=1 then orderer=38;
  if name="NLLRDTH" and level=2 then orderer=39;
  if name="NLLRDTH" and level in(. 0 1) then delete;
  if name="LIVER_TRANSPLANT" and level=0 then delete;
  if name="LIVER_TRANSPLANT" and level=1 then orderer=40;

proc sort data=accumall2;
  by JUDGE_FLAG orderer;

*** Output ***;

proc freq data=analysis;
  tables flow_num judge_flag/missing list;
  format flow_num fnratio. judge_flag judge.;
  title3 'Table N counts';

proc print data=accumall noobs;
  var orderer LEVEL name CHARALL;
  by flow_num;
  pageby flow_num;
  format flow_num fnratio.;
  title3 'Table 1';

proc print data=accumall2 noobs;
  var orderer LEVEL name CHARALL;
  by JUDGE_FLAG;
  pageby JUDGE_FLAG;
  format judge_flag judge.;
  title3 'Table 2';
```

```
proc means data=analysis n min p25 median p75 max;
  var EARLY_TO_ONSET;
  title3 'EARLY_TO_ONSET: Days between exposure and DILI recognition, median (25th, 75th percentiles)';

proc freq data=analysis;
  tables AGECAT/missing list;
  title3 'Categorical Age results';

proc print data=stdlabhx_300v(obs=50);
  var DEIDNUM LABHXN06 LABHXV06 LABHXU06;
  title3 'Sample Eosinophil Results';

proc freq data=judgment;
  tables FORM/missing list;
  title3 'FORM Results';
```